

# Dielectric screening in a spherical cavity

Christopher J Glosser and Roger C Hill

Department of Physics, Southern Illinois University, Edwardsville, IL 62026-1654

E-mail: cglosse@siue.edu, rhill@siue.edu

**Abstract.**

In this work we examine the electrostatic screening potential due to a point charge located off-centre in a spherical dielectric cavity. This potential is expanded for the case in which the dielectric constant  $\epsilon$  is large, several methods of finding the terms in the expansion are investigated, and closed-form expressions are found through third order in  $\epsilon$  along with error bounds. Finally, possible uses of these expressions in molecular dynamics simulations of isolated charged molecules is discussed.

PACS numbers: 02.30.-f, 02.70.Ns, 41.20.Cv, 87.10.-e

Submitted to: *J. Phys. A: Math. Gen.*

## 1. Introduction

Over the last generation, molecular dynamics simulations have become a vital theoretical tool in the analysis of the physical interaction of proteins. This has been driven in part by the dramatic increase in cheap computer power in the last decade, which has grown at almost an exponential rate.

This growth in computer power has resulted in a similarly dramatic increase in the size of the systems studied utilizing this technique. What began as a study of modest proteins such as myoglobin has expanded to include complicated systems such as structures embedded in cellular membranes and even a tobacco mosaic virus [1].

Due to the complexity of these systems, previously ignorable errors due to approximations in the model are likely to accumulate, resulting in inaccurate results and unstable simulations. Therefore, it is of vital importance to implement as accurate a representation as possible, particularly with respect to the long-range interactions in the model. Of these, the most problematic is the electrostatic interaction between charged elements in the simulation. In addition to generating long-range forces, the electrostatic field also polarizes the media in which the simulation is taking place, effectively creating more sources for the field in the simulation. It is this aspect of the electrostatic interaction that is the most troublesome to implement accurately while keeping computational time and expense to a minimum.

There have been numerous attempts to circumvent the electrostatics problem in molecular dynamics models. The classic way of doing this is to place the system in a periodic cell and implement Particle Mesh Ewald dynamics [2] to account for the long range fields. This model indeed handles the electrostatic problem while keeping the system size reasonable. However, it artificially imposes a crystalline structure on the system which may not be desirable for some applications.

If one wishes to investigate an isolated structure, then the options are fairly limited. A cutoff on the electrostatic interaction is usually imposed, but this effectively isolates portions of the system from one another. These models also suffer from the defect that the system in effect becomes finite in size, ignoring a large portion of the solvent. Since the solvent — which is usually water — has a large dielectric constant ( $\epsilon \approx 80$ ), it is quite polarizable. Therefore, the field generated by the solvent is very sensitive to the background field. Multipole methods [3] historically have had some success in dealing with these long range terms.

We wish to reformulate the approach to the electrostatic interaction in molecular dynamics simulations to take into account this sensitive dependence of the system on the background field. Our model should have the following properties,

- It should accurately represent the field of the solvent.
- It should be relatively inexpensive from a computational point of view.
- The potential in question should be a solution to Poisson's equation, so that it represents a physically possible charge distribution.

## 2. Green's Function for the Screening Potential

If a charge distribution is placed in a dielectric medium that is uniform and infinite in extent, the well known result is that the electric potential is reduced by a factor of  $\epsilon$ , the relative permittivity of the dielectric. This reduction is caused by an additional "screening potential" due to the polarization induced in the dielectric, which partly

cancels the original potential. The problem is more complicated when there are dielectric boundaries involved, as in the case of a charge distribution inside a cavity within a dielectric. The interaction of charges embedded in a dielectric cavity is a surprisingly complicated and rich subject in the study of classical electromagnetic theory. Even simple systems fail to have closed-form solutions for the potential. If one wishes to construct a realistic model in which charge interacts with a dielectric, then some approximation is inevitably necessary.

To begin building our model, let us assume that a dielectric of uniform relative permittivity  $\epsilon$  fills all of space except for a spherical cavity centred at the origin, and that the cavity has unit radius (i.e. all distances are measured in terms of the cavity radius). Suppose that there is a charge distribution  $\rho(\mathbf{r})$  within the cavity. The potential everywhere in space may be calculated if the Green's function  $G(\mathbf{r}; \mathbf{r}')$  is known:

$$\Phi(\mathbf{r}) = \int_{\text{cavity}} G(\mathbf{r}; \mathbf{r}') \rho(\mathbf{r}') d^3\mathbf{r}'. \quad (1)$$

The Green's function satisfies Poisson's equation as a function of  $\mathbf{r}$  with a unit point charge located at  $\mathbf{r}'$  as the source:

$$\nabla^2 G(\mathbf{r}; \mathbf{r}') = -4\pi\delta^3(\mathbf{r} - \mathbf{r}'), \quad (2)$$

and must also satisfy the proper boundary conditions on the cavity wall and at infinity. This problem is easily solved by the method of images in the limit ( $\epsilon \rightarrow \infty$ ), corresponding to a cavity in a conductor. However, no closed-form expression for  $G(\mathbf{r}; \mathbf{r}')$  is known in terms of elementary functions for the case of finite  $\epsilon$ , except for the trivial case in which the charge is located at the centre of the cavity. It is therefore necessary to representing the Green's function with an infinite series or other type of approximation.

Let us choose the positive  $z$ -axis to pass through the source point  $\mathbf{r}'$ , which we are assuming to be within the cavity so that  $r' < 1$ . The classic way of tackling this type of problem is to write the Green's function as an infinite series in orthogonal functions, which for the case at hand will be Legendre polynomials due to the azimuthal symmetry of the problem. The potential due to the unit point charge alone is

$$G_{\text{point}}(\mathbf{r}; \mathbf{r}') = \frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{\sqrt{r^2 + r'^2 - 2rr' \cos \theta}}, \quad (3)$$

which has a well known expansion in terms of Legendre Polynomials [4, 5]. For looking at boundary conditions we will be interested in the region near the cavity wall, for which  $r > r'$  and the expansion is

$$G_{\text{point}}(\mathbf{r}; \mathbf{r}') = \sum_{l=0}^{\infty} \frac{r'^l}{r^{l+1}} P_l(\cos \theta). \quad (4)$$

There is also a contribution  $G_{\text{screen}}(\mathbf{r}; \mathbf{r}')$  to the Green's function, the screening potential, due to the polarized dielectric:

$$G(\mathbf{r}; \mathbf{r}') = G_{\text{point}}(\mathbf{r}; \mathbf{r}') + G_{\text{screen}}(\mathbf{r}; \mathbf{r}'). \quad (5)$$

Because the effective polarization charge is only on the surface of the dielectric, the screening contribution to the Green's function must satisfy Laplace's equation inside and outside the cavity, be finite in each of these regions, and satisfy the proper

boundary conditions at the cavity wall  $r = 1$ . With these conditions in mind we can write

$$G_{\text{screen}}(\mathbf{r}; \mathbf{r}') = \begin{cases} \sum_{l=0}^{\infty} A_l r^l P_l(\cos \theta) & \text{for } r < 1, \\ \sum_{l=0}^{\infty} A_l r^{-(l+1)} P_l(\cos \theta) & \text{for } r > 1. \end{cases} \quad (6)$$

The coefficients  $A_l$  are the same in both sums to ensure continuity of the potential at  $r = 1$ , which is one of the boundary conditions.

To find the coefficients  $A_l$ , we enforce the other boundary condition, which is that the electric displacement be continuous across the boundary. This amounts to a condition on the radial derivative of the complete Green's function:

$$\left. \frac{\partial G}{\partial r} \right|_{r \rightarrow 1-} = \epsilon \left. \frac{\partial G}{\partial r} \right|_{r \rightarrow 1+}. \quad (7)$$

The result is

$$A_l = -\frac{(\epsilon - 1)(l + 1)}{1 + \epsilon(l + 1)} r^l. \quad (8)$$

Substituting these into (6) we find our expression for the screening Green's function to be

$$G_{\text{screen}}(\mathbf{r}; \mathbf{r}') = \begin{cases} \overline{G}(rr', \cos \theta, \epsilon) & \text{for } r < 1, \\ \frac{1}{r} \overline{G}\left(\frac{r'}{r}, \cos \theta, \epsilon\right) & \text{for } r > 1, \end{cases} \quad (9)$$

where

$$\overline{G}(x, u, \epsilon) \equiv -(\epsilon - 1) \sum_{l=0}^{\infty} \frac{l + 1}{l + \epsilon(l + 1)} x^l P_l(u) \quad (10)$$

for  $0 \leq x \leq 1$  and  $-1 \leq u \leq 1$ . This series solution is well known in the literature [6, 7, 8]. We shall refer to the function  $\overline{G}(x, u, \epsilon)$  as the “screening function.”

While Equation (10) provides a perfectly legitimate expression for the screening function, it is fairly limited to its usefulness in numerical simulations of charges within the cavity. The rate of convergence depends on where the charge is within the cavity, and is rather slow unless the charge is close to the origin. However, we can look for methods of summing the series with hopes of obtaining an expression that may be more useful in calculations. The coefficient within the sum in (10) can be written as

$$\frac{l + 1}{l + \epsilon(l + 1)} = \frac{1}{\epsilon + 1} + \frac{1}{(\epsilon + 1)^2} \cdot \frac{1}{l + \epsilon/(\epsilon + 1)}, \quad (11)$$

separating the sum into two parts:

$$\overline{G}(x, u, \epsilon) = -\frac{\epsilon - 1}{\epsilon + 1} \left[ \sum_{l=0}^{\infty} x^l P_l(u) + \frac{1}{\epsilon + 1} \sum_{l=0}^{\infty} \frac{1}{l + \epsilon/(\epsilon + 1)} x^l P_l(u) \right]. \quad (12)$$

The first is just the generating function for the Legendre polynomials [9],

$$\sum_{l=0}^{\infty} x^l P_l(u) = \frac{1}{\sqrt{1 - 2ux + x^2}}. \quad (13)$$

The other sum can be evaluated in several ways, one of the simplest of which is to start with the generating function again,

$$\frac{1}{\sqrt{1-2ut+t^2}} = \sum_{l=0}^{\infty} t^l P_l(u), \quad (14)$$

multiply both sides by  $t^{\epsilon/(\epsilon+1)-1}$ , and integrate:

$$\begin{aligned} \int_0^x \frac{t^{\epsilon/(\epsilon+1)-1}}{\sqrt{t^2-2tu+1}} dt &= \int_0^x t^{\epsilon/(\epsilon+1)-1} \sum_{l=0}^{\infty} t^l P_l(u) dt \\ &= \sum_{l=0}^{\infty} \frac{x^{l+\epsilon/(\epsilon+1)}}{l+\epsilon/(\epsilon+1)} P_l(u). \end{aligned} \quad (15)$$

This allows us to express the Green's function in terms of an integral:<sup>‡</sup>

$$\overline{G}(x, u, \epsilon) = -\frac{\epsilon-1}{\epsilon+1} \left[ \frac{1}{\sqrt{1-2ux+x^2}} + \frac{x^{-\epsilon/(\epsilon+1)}}{1+\epsilon} \int_0^x \frac{t^{-1/(\epsilon+1)} dt}{\sqrt{1-2ut+t^2}} \right]. \quad (16)$$

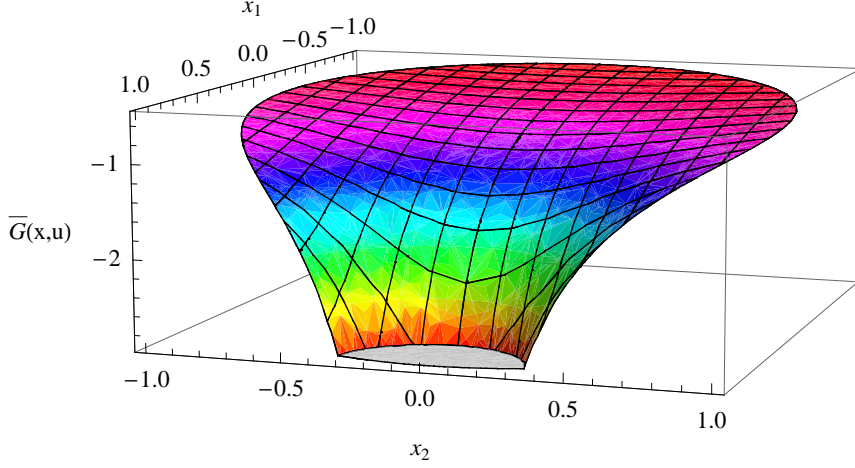
Expression (16) lends itself to the following physical interpretation: First of all, in the limit  $\epsilon \rightarrow \infty$ , where the dielectric becomes a conductor, only the first term in the square brackets survives. For the solution inside the cavity, this corresponds to the familiar image point charge outside the cavity, and for the solution outside the cavity, this corresponds to putting an image charge at the location of the original charge, neutralizing the original charge and “grounding” the conductor. Now for finite  $\epsilon$ , the image charge is reduced in magnitude by the factor  $(\epsilon-1)/(\epsilon+1)$  so that there is a nonzero potential outside the cavity, but this point charge is not enough to satisfy the boundary conditions; an additional *line* image charge is needed, stretching from the image point charge to infinity (for the inner solution) or to the origin (for the outer solution).

Figure 1 shows a plot of the screening function as a function of the variables  $x_1 = x \cos \theta$  and  $x_2 = x \sin \theta$  in the circular region defined by  $0 \leq x \leq 1$ . For points inside the cavity,  $x = rr'$  as indicated in (9), so  $x_1$  and  $x_2$  are proportional to the spatial coordinates parallel and perpendicular to the  $z$ -axis, respectively. For a given location  $r'$  of the source charge, the cavity boundary corresponds to  $x = r'$  which is at most one, so a plot of the potential inside the cavity may be visualized by truncating the plot in figure 1 to a smaller circular region of radius  $r'$ , and then expanding the plot to fill a region of radius one. The screening function becomes infinite at the point  $(x=1, u=1)$ , corresponding to the image point charge referred to in the previous paragraph, but this singularity only shows up in the physical solution when  $r' = 1$ , i.e. when the source charge is at the cavity wall.

### 3. Expansion for Large Values of the Dielectric Constant

Equation (16) provides a concise, exact expression for the screening function, which can be evaluated numerically and also in terms of Appell hypergeometric functions [10]. However, it is of limited use in actual simulations of molecules in which great numbers of these expressions would need to be evaluated. It is therefore useful to look for approximations allowing the use of simpler functions. One such approximation is to recognise that, as we have mentioned, the dielectric constant of water is quite large, so that an expansion good for large values of  $\epsilon$  would be useful.

<sup>‡</sup> This equation is equivalent to the one appearing in references [7, 8].



**Figure 1.** A plot of the screening function for  $\epsilon = 80$ , produced by numerically evaluating and plotting (16) using the standard packages in *Mathematica*. The axis coordinates in this and subsequent surface plots were chosen to be  $x_1 = xu = x \cos \theta$  and  $x_2 = \pm x\sqrt{1-u^2} = \pm x \sin \theta$ .

Based on what we have already derived here, there are two ways in which we can obtain a series expansion good for large  $\epsilon$ . One way is to go back to definition (10) of  $G_n(x, u)$  and write it in the form

$$\begin{aligned} G_n(x, u) &= -\frac{\epsilon-1}{\epsilon+1} \sum_{l=0}^{\infty} \left[ 1 - \frac{1}{(\epsilon+1)(l+1)} \right]^{-1} x^l P_l(u) \\ &= -\frac{\epsilon-1}{\epsilon+1} \sum_{l=0}^{\infty} \sum_{n=0}^{\infty} \frac{1}{(\epsilon+1)^n (l+1)^n} x^l P_l(u). \end{aligned} \quad (17)$$

Reversing the order of summation gives us a series in powers of  $1/(\epsilon+1)^n$ :

$$\overline{G}(x, u, \epsilon) = -\frac{\epsilon-1}{\epsilon+1} \left( G_0(x, u) + \frac{G_1(x, u)}{\epsilon+1} + \frac{G_2(x, u)}{(\epsilon+1)^2} + \frac{G_3(x, u)}{(\epsilon+1)^3} + \dots \right), \quad (18)$$

where

$$G_n(x, u) = \sum_{l=0}^{\infty} \frac{x^l}{(l+1)^n} P_l(u). \quad (19)$$

The other way to get a series is to expand the integral expression (16), noting that

$$x^{-\epsilon/(\epsilon+1)} t^{-1/(\epsilon+1)} = \frac{1}{x} \exp \frac{\ln(x/t)}{\epsilon+1} = \frac{1}{x} \sum_{n=1}^{\infty} \frac{\ln^{n-1}(x/t)}{(n-1)! (\epsilon+1)^{n-1}}. \quad (20)$$

This gives the series in (18) again, where this time the coefficients are given by

$$G_0(x, u) = \frac{1}{\sqrt{1-2ux+x^2}} \quad (21)$$

and

$$G_n(x, u) = \frac{1}{(n-1)!} \cdot \frac{1}{x} \int_0^x \frac{\ln^{n-1}(x/t)}{\sqrt{1-2ut+t^2}} dt \quad \text{for } n \geq 1. \quad (22)$$

Expression (21) is clearly equivalent to series expression (19) for  $n = 0$ , since the former is the generating function for the Legendre polynomials. It is also possible to show directly that series expression (19) and integral expression (22) are equivalent for  $n \geq 0$ . Our purpose in deriving these two expressions for  $G_n(x, u)$  is that they complement each other: The series is useful for deriving general properties of these coefficients, while the integral is useful for doing actual calculations of them.

To get an idea of what sort of functions may be involved in finding the coefficients  $G_n(x, u)$ , let us use the series expression to evaluate them in the special case  $u = \pm 1$ , corresponding to points on the  $z$ -axis. Since  $P_l(\pm 1) = (\pm 1)^l$ , equation (19) gives

$$G_n(x, \pm 1) = \sum_{l=0}^{\infty} (\pm 1)^l \frac{x^l}{(l+1)^n} = \pm \frac{1}{x} \text{Li}_n[\pm x], \quad (23)$$

where  $\text{Li}_n$  is the polylogarithm function [11, 12, 13]:

$$\text{Li}_n[x] \equiv \sum_{k=1}^{\infty} \frac{x^k}{k^n}. \quad (24)$$

As can be easily seen from this definition, the polylogarithms satisfy a recursion relation,

$$\text{Li}_n[x] = \int_0^x \frac{\text{Li}_{n-1}[t]}{t} dt. \quad (25)$$

The polylogarithm for  $n = 0$  is easily found from definition (24):

$$\text{Li}_0[x] = \frac{x}{1-x}, \quad (26)$$

and those for  $n = 1, 2, 3, \dots$  may be obtained from this function by successive integrations. As a consequence,

$$\text{Li}_1[x] = -\ln(1-x), \quad (27)$$

while the dilogarithm  $\text{Li}_2[x]$ , the trilogarithm  $\text{Li}_3[x]$ , and higher-order polylogarithms cannot be expressed in terms of elementary functions. For any positive integer  $n$ ,  $\text{Li}_n[x]$  has a real value for  $-\infty < x \leq 1$ , but has a branch cut in the complex plane running along the positive real axis from  $x = 1$  to  $\infty$ , across which the function has a discontinuous imaginary part.

Like the polylogarithms, the coefficients  $G_n(x, u)$  also have a similar recursion relation which follows immediately from series expression (19):

$$G_n(x, u) = \frac{1}{x} \int_0^x G_{n-1}(t, u) dt, \quad (28)$$

and in fact the recursion relation for  $xG_n(x, u)$  is the same as for  $\text{Li}_n[x]$ . Equation (28) serves as an alternate way to calculate  $G_n(x, u)$  by starting with  $G_0(x, u)$  as given by (21).

To calculate  $G_n(x, u)$  for specific values of  $n$ , we can either use integral expression (22) or recursion relation (28); both methods seem to lead to about the same degree of complexity. In both methods we have found it useful to make the change of variable

$$w = \frac{1+t-\sqrt{1-2ut+t^2}}{1+u}, \quad \text{or} \quad t = \frac{w}{1-w} \left(1 - \frac{1+u}{2} w\right). \quad (29)$$

In addition, we define the following variables for the sake of convenience,

$$p \equiv \frac{1+x-\sqrt{1-2ux+x^2}}{1+u}, \quad q \equiv \frac{1-u}{1+u}. \quad (30)$$

For physical values of  $x$  and  $u$ , the quantities  $p$  and  $q$  lie in the range  $0 \leq p \leq 1$  and  $0 \leq q < \infty$ .<sup>§</sup> The first-order coefficient  $G_1(x, u)$  is then easily evaluated using either the integral expression or the recursion relation; both methods lead to the same integral:

$$\begin{aligned} G_1(x, u) &= \frac{1}{x} \int_0^x \frac{dt}{\sqrt{1-2ut+t^2}} = \frac{1}{x} \int_0^p \frac{dw}{1-w} \\ &= -\frac{1}{x} \ln(1-p) = \frac{1}{x} \text{Li}_1[p], \end{aligned} \quad (31)$$

The higher-order coefficients involve higher-order polylogarithms and become increasingly complicated. A pattern that emerges is that for  $n \geq 1$ ,  $xG_n(x, u)$  can be expressed entirely in terms of logarithms and polylogarithms of rational functions of  $p$  and  $q$ . Using transformation (29) on integral (22), along with (30) to eliminate  $x$  in the integrand, we obtain

$$xG_n(x, u) = \frac{1}{(n-1)!} \int_0^p \ln^{n-1} \left[ \frac{b(1-w)}{w(1+q-w)} \right] \frac{dw}{1-w}, \quad (32)$$

where  $b = p(1+q-p)/(1-p)$ . Also, from (28) we obtain a transformed recursion relation,

$$xG_n(x, u) = \frac{1}{x} \int_0^p tG_{n-1}(t, u) \left( \frac{1}{w} + \frac{1}{1-w} - \frac{1}{1+q-w} \right) dw, \quad (33)$$

where  $tG_{n-1}(t, u)$  is assumed to be written in terms of  $w$  and  $q$ . Armed with either of these equations we can evaluate the second-order coefficient without much difficulty. For example, using (33) we have

$$xG_2(x, u) = - \int_0^p \frac{\ln(1-w)}{w} dw - \int_0^p \frac{\ln(1-w)}{1-w} dw + \int_0^p \frac{\ln(1-w)}{1+q-w} dw. \quad (34)$$

The first integral is just  $\text{Li}_2[p]$  and the second one is easily evaluated as  $\frac{1}{2} \ln^2(1-p)$ . The third integral can be found by making the change of variable  $w' = q/(1+q-w)$ ; the result is  $-\text{Li}_2[w'] - \ln(w'/q)$  evaluated at the endpoints. The final result is

$$\begin{aligned} G_2(x, u) &= \frac{1}{x} \left\{ \text{Li}_2[p] + \text{Li}_2 \left[ \frac{q}{1+q} \right] - \text{Li}_2 \left[ \frac{q}{1-p+q} \right] \right. \\ &\quad \left. + \frac{1}{2} \ln^2(1-p) + \frac{1}{2} \ln^2(1+q) - \frac{1}{2} \ln^2(1-p+q) \right\}. \end{aligned} \quad (35)$$

The third-order coefficient is considerably more complicated, whether we use (32) or (33). We evaluated it using (32), enlisting the aid of *Mathematica* [14] to find and manipulate the large number of terms and help simplify the expression. In the process we also made use of a number of dilogarithm and trilogarithm identities [12, 13]. Our result is

$$\begin{aligned} G_3(x, u) &= \frac{1}{x} \left\{ 2 \text{Li}_3[p] - \text{Li}_3 \left[ \frac{p}{1+q} \right] + \text{Li}_3[1-p] - \text{Li}_3 \left[ \frac{1-p+q}{1+q} \right] \right. \\ &\quad \left. - \text{Li}_3 \left[ \frac{1-p}{1-p+q} \right] + \text{Li}_3 \left[ \frac{1}{1+q} \right] - \text{Li}_3 \left[ \frac{q}{1+q} \right] + \text{Li}_3 \left[ \frac{q}{1-p+q} \right] \right\} \end{aligned}$$

<sup>§</sup> The ranges of these variables are interconnected by the fact that, for a given value of  $q$ , the maximum value of  $p$  (corresponding to  $x = 1$ ) is  $1+q - \sqrt{q(1+q)}$ , which is 1 in the limit  $u \rightarrow \pm 1$  but less than 1 otherwise.



$$\begin{aligned}
& -\text{Li}_3 \left[ \frac{pq}{1-p+q} \right] + \ln \frac{(1-p+q)^2}{(1-p)(1+q)^2} \text{Li}_2 \left[ \frac{q}{1+q} \right] \\
& + \frac{1}{6} \ln^3 \frac{1-p+q}{(1-p)(1+q)} - \frac{1}{2} \ln \frac{pq}{1+q} \ln^2 \frac{1-p+q}{(1-p)(1+q)} \\
& - \left[ \ln p \ln(1-p) + \ln(1+q) \ln \frac{q}{1+q} \right] \ln \frac{1-p+q}{(1-p)(1+q)} \Big\}.
\end{aligned} \tag{36}$$

One of the complications in writing down the coefficients past first order is that there are many different-looking ways to express each  $G_n(x, u)$ , due to the numerous identities satisfied by the polylogarithms. In an attempt to be somewhat systematic in our expressions, and also to facilitate the study of their behaviour as well as repeated calculations that would occur in plotting or numerical simulations, we have used these identities where necessary to express the results in a form where the arguments of all polylogarithms are between 0 and 1 inclusive.

As a check on these results, we can look at their behaviour in the limit  $u \rightarrow 1$ , which corresponds to

$$p \rightarrow x, \quad q \rightarrow 0. \tag{37}$$

In (35), the first term in the curly brackets becomes  $\text{Li}_2[x]$  while the other terms either become zero or cancel in pairs, so that  $G_2(x, 1) = \text{Li}_2[x]/x$  as expected from (23). In (36), the first two terms give  $\text{Li}_3[x]$  while the rest becomes zero, giving the expected result  $G_3(x, 1) = \text{Li}_3[x]/x$ . The behaviour in the limit  $u \rightarrow -1$ , while correct, is less straightforward; it corresponds to

$$p \rightarrow \frac{x}{1+x}, \quad q \rightarrow \infty, \tag{38}$$

and in this limit we obtain

$$G_2(x, -1) = \frac{1}{x} \left\{ \text{Li}_2 \left[ \frac{x}{1+x} \right] + \frac{1}{2} \ln^2(1+x) \right\}, \tag{39}$$

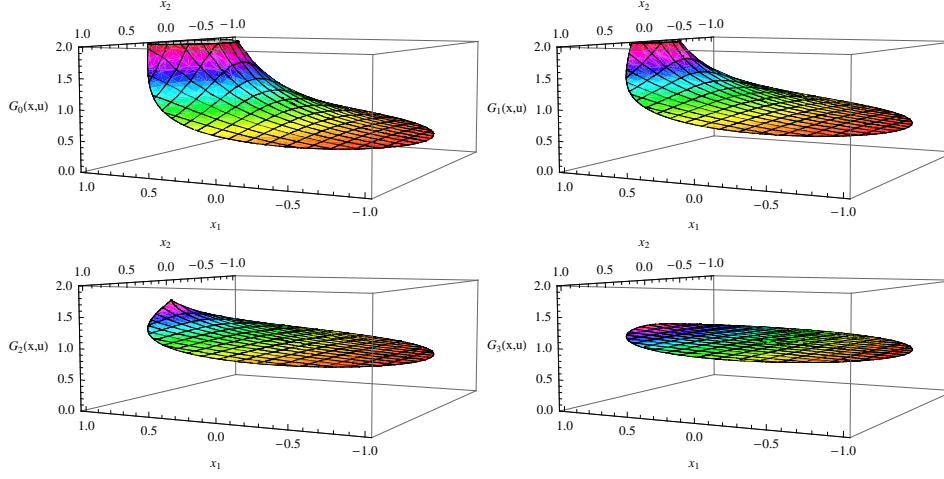
which reduces to the expected  $-\text{Li}_2[-x]/x$  by way of one of the dilog identities. The behaviour of the third-order coefficient is similar.

It should be pointed out that despite the factor of  $1/x$  occurring in these expressions,  $G_n(x, u)$  is not singular at  $x = 0$ , because the rest of the expression becomes zero there. In fact,

$$G_n(0, u) = 1, \tag{40}$$

as can be seen from series expression (19).

Figure 2 shows three-dimensional plots of  $G_0(x, \cos \theta)$  through  $G_3(x, \cos \theta)$  as functions of  $x_1 = x \cos \theta$  and  $x_2 = x \sin \theta$ , similarly to figure 1. The zeroth-order coefficient has an inverse first-power singularity at the point  $(x = 1, u = 1)$  corresponding to the image point charge as per the discussion following equation (16), and the first-order coefficient has a logarithmic infinity at that point, while the rest of the coefficients are finite everywhere in the region. Despite the increasing complexity of the expressions as the order becomes higher, their actual behaviour becomes increasingly simple; the plot of the third-order term is comparatively flat. We can in fact obtain bounds on  $G_n(x, u)$  in general by noting from integral expression (22) that for any given  $x$ , the integral will have a maximum value when  $u = 1$  and a minimum value when  $u = -1$ . It therefore follows from (23) that for given  $x$ , the maximum value is  $\text{Li}_n[x]/x$  and the minimum value is  $-\text{Li}_n[-x]/x$ . Since for  $0 \leq x \leq 1$ ,  $\text{Li}_n[x]/x$  and  $-\text{Li}_n[-x]/x$  are strictly increasing and decreasing functions



**Figure 2.** The coefficient functions  $G_0$  through  $G_3$ .

of  $x$  respectively,|| it follows that (for  $n \geq 1$ ) the maximum value over all  $x$  and  $u$  is  $\text{Li}_n[1] = \zeta(n)$ , and the minimum value is  $-\text{Li}_n[-1] = (1 - 2^{-n+1})\zeta(n)$ , where  $\zeta(n)$  is the Riemann zeta function:

$$\zeta(n) \equiv \frac{1}{1^n} + \frac{1}{2^n} + \frac{1}{3^n} + \dots \quad (41)$$

It follows that

$$(1 - 2^{-n+1})\zeta(n) \leq G_n(x, u) \leq \zeta(n) \quad (42)$$

for all  $x$  and  $u$  in the physical region, and for  $n \geq 1$ . The  $n = 0$  case is easily treated; the lower bound is  $G_0(1, -1) = \frac{1}{2}$ . The results are numerically,

$$\begin{aligned} 0.500000 &\leq G_0(x, u) < \infty \\ 0.693147 &\leq G_1(x, u) < \infty \\ 0.822467 &\leq G_2(x, u) \leq 1.644934 \\ 0.901543 &\leq G_3(x, u) \leq 1.202057. \end{aligned} \quad (43)$$

As  $n$  becomes large, the lower and upper bounds both approach 1, leading to an increasingly flat plot. The coefficient  $G_n(x, u)$  could then be approximated by a simple polynomial, or even a constant, greatly reducing time in computationally intensive problems. However, for problems occurring in practice with large values of the dielectric constant, it may not be necessary to keep very many terms anyway. We will look at the errors due to truncating the series in section 5.

#### 4. Electric Fields

While the potential is useful for calculating the potential energy of the system, one also needs to be able to calculate the force of the various objects upon one another so that the system may evolve from one time step to the next. This in turn requires knowing the electric field due to the charges and the dielectric. Corresponding to the

|| This can be seen by looking at their derivatives, using the power series in (23).

Green's function  $G(\mathbf{r}; \mathbf{r}')$ , which is the potential at  $\mathbf{r}$  due to a unit point charge at  $\mathbf{r}'$  in the presence of the dielectric, let us define  $\mathbf{E}(\mathbf{r}; \mathbf{r}')$  be the electric field at  $\mathbf{r}$  due to a unit point charge at  $\mathbf{r}'$  in the presence of the dielectric. The electric field is determined from the Green's function by

$$\mathbf{E}(\mathbf{r}; \mathbf{r}') = -\nabla_{\mathbf{r}} G(\mathbf{r}; \mathbf{r}'), \quad (44)$$

where  $\nabla_{\mathbf{r}}$  means the gradient with respect to the vector  $\mathbf{r}$ , keeping  $\mathbf{r}'$  constant. Like the potential, the electric field separates into the field to the point charge alone plus the "screening field" due to the polarized dielectric:

$$\mathbf{E}(\mathbf{r}; \mathbf{r}') = \mathbf{E}_{\text{point}}(\mathbf{r}; \mathbf{r}') + \mathbf{E}_{\text{screen}}(\mathbf{r}; \mathbf{r}'). \quad (45)$$

The field  $\mathbf{E}_{\text{point}}(\mathbf{r}; \mathbf{r}')$  is the familiar Coulomb field,

$$\mathbf{E}_{\text{point}}(\mathbf{r}; \mathbf{r}') = \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} = \frac{(r - r' \cos \theta) \mathbf{e}_r + (r' \sin \theta) \mathbf{e}_\theta}{(r^2 + r'^2 - 2rr' \cos \theta)^{3/2}}, \quad (46)$$

where  $\mathbf{e}_r$  and  $\mathbf{e}_\theta$  are unit vectors in the direction of increasing  $r$  and  $\theta$ . In (9) we expressed the screening part of the Green's function inside and outside of the cavity in terms of a single function  $\overline{G}(x, u, \epsilon)$ , which can be thought of as a function of a three-dimensional coordinate  $\mathbf{x}$  with spherical coordinates  $(x, \theta, \phi)$ . Let us define  $\overline{\mathbf{E}}(x, u, \epsilon)$  to be the negative of the gradient of this function:

$$\begin{aligned} \overline{\mathbf{E}}(x, u, \epsilon) &= -\nabla_{\mathbf{x}} \overline{G}(x, u, \epsilon) = -\left( \mathbf{e}_x \frac{\partial}{\partial x} + \mathbf{e}_\theta \frac{1}{x} \frac{\partial}{\partial \theta} \right) \overline{G}(x, \cos \theta, \epsilon) \\ &= \left( -\mathbf{e}_x \frac{\partial}{\partial x} + \mathbf{e}_\theta \frac{\sqrt{1-u^2}}{x} \frac{\partial}{\partial u} \right) \overline{G}(x, u, \epsilon). \end{aligned} \quad (47)$$

Then from (9), the screening field can be written as

$$\mathbf{E}_{\text{screen}}(\mathbf{r}; \mathbf{r}') = \begin{cases} r' \overline{\mathbf{E}}(r' r, \cos \theta, \epsilon) & \text{for } r < 1, \\ \frac{r'}{r^3} \overline{\mathbf{E}}\left(\frac{r'}{r}, \cos \theta, \epsilon\right) + \frac{1}{r^2} \overline{G}\left(\frac{r'}{r}, \cos \theta, \epsilon\right) \mathbf{e}_r & \text{for } r > 1. \end{cases} \quad (48)$$

The expansion of  $\overline{G}(x, u, \epsilon)$  in powers of  $1/(1+\epsilon)$  leads to a corresponding expansion of  $\overline{\mathbf{E}}(x, u, \epsilon)$ :

$$\overline{\mathbf{E}}(x, u, \epsilon) = -\frac{\epsilon - 1}{\epsilon + 1} \left( \mathbf{E}_0(x, u) + \frac{\mathbf{E}_1(x, u)}{\epsilon + 1} + \frac{\mathbf{E}_2(x, u)}{(\epsilon + 1)^2} + \frac{\mathbf{E}_3(x, u)}{(\epsilon + 1)^3} + \dots \right), \quad (49)$$

where

$$\mathbf{E}_n(x, u) = -\nabla_{\mathbf{x}} G_n(x, u), \quad (50)$$

The zero-order term is found in a straightforward manner, and is just the Coulomb electric field due to a unit point charge located on the  $x$ -axis at unit distance from the origin:

$$\mathbf{E}_0(x, u) = \frac{(x - u) \mathbf{e}_x + \sqrt{1 - u^2} \mathbf{e}_\theta}{(1 - 2ux + x^2)^{3/2}}. \quad (51)$$

For the higher orders, we can avoid having to take some of the derivatives of  $G_n(x, u)$  explicitly by noticing that if we differentiate both sides of recursion relation (28) with respect to  $x$ , we obtain

$$\frac{\partial G_n}{\partial x} = \frac{1}{x} (G_{n-1} - G_n) \quad \text{for } n \geq 1. \quad (52)$$

The derivative with respect to  $u$  is not so easily found, but even it can be simplified. For  $n \geq 1$ , we found that our expressions for  $G_n(x, u)$  were of the form  $1/x$  times a function consisting of logarithms and polylogarithms of  $p$  and  $q$ . Therefore, for  $n \geq 1$  it is convenient to write

$$\frac{\partial G_n}{\partial u} = \frac{1}{x} \frac{\partial(xG_n)}{\partial u} = \frac{1}{x} \left[ \frac{\partial p}{\partial u} \frac{\partial(xG_n)}{\partial p} + \frac{\partial q}{\partial u} \frac{\partial(xG_n)}{\partial q} \right]. \quad (53)$$

On the other hand we also know that, again from the recursion relation,

$$G_{n-1} = \frac{\partial(xG_n)}{\partial x} = \frac{\partial p}{\partial x} \frac{\partial(xG_n)}{\partial p} + \frac{\partial q}{\partial x} \frac{\partial(xG_n)}{\partial q}. \quad (54)$$

From definitions (30) of  $p$  and  $q$ , the partial derivatives of them with respect to  $x$  and  $p$  are

$$\begin{aligned} \frac{\partial p}{\partial x} &= \frac{(1-p)^2(1+q)}{(1-p)^2 + q}, & \frac{\partial p}{\partial u} &= \frac{p^2(1-p)(1+q)}{2[(1-p)^2 + q]}, \\ \frac{\partial q}{\partial x} &= 0, & \frac{\partial q}{\partial u} &= -\frac{(1+q)^2}{2}. \end{aligned} \quad (55)$$

Since  $\partial q/\partial x = 0$ , equation (54) allows us to express  $\partial(xG_n)/\partial x$  in terms of  $G_{n-1}$ , which can then be used in (53), giving

$$\begin{aligned} \frac{\partial G_n}{\partial u} &= \frac{1}{x} \left[ \frac{\partial p/\partial u}{\partial p/\partial x} G_{n-1} + \frac{\partial q}{\partial u} \frac{\partial(xG_n)}{\partial q} \right] \\ &= \frac{1}{x} \left[ \frac{p^2}{2(1-p)} G_{n-1} - \frac{(1+q)^2}{2} \frac{\partial(xG_n)}{\partial q} \right] \end{aligned} \quad (56)$$

Using these results in (47) along with the fact that  $\sqrt{1-u^2} = 2\sqrt{q}/(1+q)$ , we obtain for the  $n^{\text{th}}$ -order coefficient in the electric field

$$\mathbf{E}_n(x, u) = \frac{1}{x} (G_n - G_{n-1}) \mathbf{e}_x + \frac{\sqrt{q}}{x^2} \left[ \frac{p^2}{(1-p)(1+q)} G_{n-1} - (1+q) \frac{\partial(xG_n)}{\partial q} \right] \mathbf{e}_\theta. \quad (57)$$

so that we only need to take derivatives with respect to  $q$ . It should be noted that, like  $G_n(x, u)$ ,  $\mathbf{E}_n(x, u)$  is not singular at  $x = 0$  despite the inverse powers of  $x$ , as they multiply factors which become zero there. In simplifying expressions for the potential and electric field it is sometimes convenient to use the expressions for  $x$  and  $G_0$  in terms of  $p$  and  $q$ :

$$x = \frac{p(1-p+q)}{(1-p)(1+q)}, \quad (58)$$

$$G_0(x, u) = \frac{1}{\sqrt{1-2ux+x^2}} = \frac{(1-p)(1+q)}{(1-p)^2 + q}. \quad (59)$$

As an example in finding a higher-order coefficient in the electric field, let us evaluate  $\mathbf{E}_1(x, u)$ . From (31),  $xG_1 = -\ln(1-p)$  which is independent of  $q$ , so no derivatives have to be taken at all in evaluating (57). After some simplification we obtain

$$\mathbf{E}_1(x, u) = -\frac{1}{x^2} \left[ \frac{p(1-p+q)}{(1-p)^2 + q} + \ln(1-p) \right] \mathbf{e}_x + \frac{1}{x^2} \frac{p^2 \sqrt{q}}{(1-p)^2 + q} \mathbf{e}_\theta. \quad (60)$$

The  $(1-p)^2 + q$  in the denominator will be recognized as the inverse first-power singularity characteristic of  $G_0(x, u)$ . As with the potential, the electric field is finite at  $x = 0$  despite the inverse powers of  $x$ .

## 5. Analysis of Errors

In implementing our series expansion (18) for any particular problem, we need to know when we can truncate the series for a given desired degree of accuracy (particularly considering how increasingly complicated the series coefficients become with each subsequent order). Suppose that we have evaluated the series from  $n = 0$  to  $n = N$ . The error  $e_N$  due to ignoring the terms from  $N + 1$  to infinity is

$$e_N = \frac{\epsilon - 1}{\epsilon + 1} \sum_{n=N+1}^{\infty} \frac{G_n(x, u)}{(\epsilon + 1)^n}. \quad (61)$$

The terms in this series are all positive, and the maximum value of  $G_n(x, u)$  is  $\zeta(n)$  from (42), so

$$e_N \leq \frac{\epsilon - 1}{\epsilon + 1} \sum_{n=N+1}^{\infty} \frac{\zeta(n)}{(\epsilon + 1)^n} \quad (62)$$

This sum can be evaluated numerically for any given  $\epsilon$  and  $N$ . For example, using  $\epsilon = 80$  we find

$$\begin{aligned} e_1 &\leq 2.468 \times 10^{-4} \\ e_2 &\leq 2.231 \times 10^{-6} \end{aligned} \quad (63)$$

$$e_3 \leq 2.482 \times 10^{-8}. \quad (64)$$

We can also find an upper bound for the *fractional* error  $e_N/|\overline{G}(x, u, \epsilon)|$  by noting that the true value of  $|\overline{G}(x, u, \epsilon)|$  is greater than the  $N^{\text{th}}$ -order calculated value, and therefore the fractional error will be less than that found by dividing the error by the calculated value. Also, a lower bound over all  $x$  and  $u$  for the calculated value can be found using the lower bounds for  $G_n(x, u)$  in (43). For example, suppose that we have kept only the  $n = 0$  and  $n = 1$  terms, again assuming  $\epsilon = 80$ . The maximum error is  $e_1$  in (63), while the minimum value of  $|\overline{G}(x, u, \epsilon)|$  is at least

$$|\overline{G}(x, u, \epsilon)|_{\min} = \frac{79}{81} \left( 0.5000 + \frac{0.6931}{80} \right) = 0.4961 \quad (65)$$

and so an upper bound for the fractional error for any  $x$  and  $u$  is  $2.468 \times 10^{-4} / 0.4961 = 4.974 \times 10^{-4}$ . Similarly, for a calculation through second order the fractional error will be less than  $4.496 \times 10^{-6}$ , and for a calculation through third order, less than  $5.001 \times 10^{-8}$ . Of course, the fractional errors will be much smaller than these bounds in regions where  $\overline{G}$  is large, even though the error itself do not vary much over the whole region.

## 6. Use in Molecular Dynamics Simulations

Having obtained our expression for the potential to third order in  $(1 + \epsilon)^{-1}$ , we now turn to the question of feasibility. That is, to what degree is this calculation useful to those doing large-scale simulations of biological systems?

A cursory objection may be that the computation of special functions as required by this potential for each of the charged objects in a simulation would be intractable, and that it would lead to an inexorable increase in computation time. While certainly true, the problem is not as severe as it first seems.

First, the computation of polylogarithms is very well understood. The Taylor series converges rapidly for arguments in the range  $0 < x < 1$ . Moreover, a partial fraction technique exists [15] that converges extremely rapidly. For instance, an evaluation of  $\text{Li}_2 \left[ \frac{1}{2} \right]$  converges to within  $1 \times 10^{-7}$  of the actual value utilizing only 10 divisions.

Second, if the system is of a suitable size, some fields can be calculated using the total charges of extended structures, such as the residues comprising the proteins, rather than the constituent atoms. Since each term in the potential is separately a solution to Laplace's equation, they each comprise a complete solution. For instance, one could approach a large system by calculating the atomic charges exactly to  $O[(1 + \epsilon)^{-2}]$ , while using the  $G_2(x, u)$  term to calculate screening by using total charge of the protein residues and treating them as point charges.

## 7. Summary

For the potential of a point charge in a spherical dielectric cavity, the authors have computed integral expressions that lead to a calculable expansion in inverse powers of the relative permittivity. The main feature of this expansion is that the truncated series yields a potential that is accurate over the physical region while remaining a solution to Poisson's equation at each order. For the case of water, truncating the series to second order leads to a fractional error of no more than  $4.5 \times 10^{-6}$ , and truncating the series to third order leads to a fractional error of no more than  $5.0 \times 10^{-8}$ .

While much of the theoretical groundwork in this article is complete, the feasibility of this model in an actual simulation environment has not been studied. The polylogarithms introduced to describe the series coefficients  $G_n(x, u)$  may be computed rapidly and precisely using the partial fraction technique, but it remains to be seen whether the accuracy gains surmount the time lag introduced by having additional terms in the potential. A systematic study of this would need to be done by altering one of the existing molecular dynamics packages such as *CHARMM* [16] or *NAMD* [17].

## Appendix A. Contour Representations

In addition to the methods that we have presented for finding the screening function and the coefficients in the expansion for large  $\epsilon$ , there are other methods using complex-variable techniques. For example, the second sum in (12) can be found by looking at the special case  $u = 1$ , which can be expressed as a hypergeometric function [9]:

$$\frac{1}{\epsilon + 1} \sum_{l=0}^{\infty} \frac{1}{l + \frac{\epsilon}{\epsilon + 1}} x^l = \frac{1}{\epsilon} {}_2F_1 \left[ \frac{\epsilon}{\epsilon + 1}, 1; 1 + \frac{\epsilon}{\epsilon + 1}; x \right]. \quad (\text{A.1})$$

The Legendre polynomials can then be brought in by using a contour-integral expression for them,

$$P_l(u) = \frac{1}{2\pi i} \oint \frac{z^{-l}}{\sqrt{1 - 2uz + z^2}} \frac{dz}{z}, \quad (\text{A.2})$$

where the integration is counterclockwise along any path that encircles the origin but does not encircle the branch points of the square root (e.g. a circle of radius less than

1). The sum with the Legendre polynomials can then be expressed in terms of a contour integral:

$$\begin{aligned} \sum_{l=0}^{\infty} \frac{x^l}{l + \frac{\epsilon}{\epsilon+1}} P_l(u) &= \frac{1}{2\pi i} \oint \sum_{l=0}^{\infty} \frac{1}{l + \frac{\epsilon}{\epsilon+1}} \frac{(z/x)^{-l}}{\sqrt{1-2uz+z^2}} \frac{dz}{z}, \\ &= \frac{1+\epsilon}{\epsilon} \frac{1}{2\pi i} \oint \frac{{}_2F_1\left[\frac{\epsilon}{\epsilon+1}, 1; 1 + \frac{\epsilon}{\epsilon+1}; \frac{x}{z}\right]}{\sqrt{1-2uz+z^2}} \frac{dz}{z}. \end{aligned} \quad (\text{A.3})$$

The screening function then becomes

$$\bar{G}(x, u, \epsilon) = -\frac{\epsilon-1}{\epsilon+1} \left[ \frac{1}{\sqrt{1-2ux+x^2}} + \frac{1}{\epsilon} \frac{1}{2\pi i} \oint \frac{{}_2F_1\left[\frac{\epsilon}{\epsilon+1}, 1; 1 + \frac{\epsilon}{\epsilon+1}; \frac{x}{z}\right]}{\sqrt{1-2uz+z^2}} \frac{dz}{z} \right]. \quad (\text{A.4})$$

A similar technique can be used in expressing the  $n^{\text{th}}$ -order coefficient  $G_n(x, u)$ , using series expression (23) for the case  $u = 1$  in combination with contour integral (A.2), giving

$$G_n(x, u) = \frac{1}{x} \frac{1}{2\pi i} \oint \frac{\text{Li}_n\left[\frac{x}{z}\right]}{\sqrt{1-2zu+z^2}} dz. \quad (\text{A.5})$$

All of the expressions for the  $G_n$  listed in the main body of the text may be obtained by directly evaluating this contour expression. In particular, it is easy to derive the recursion relation (28) by direct differentiation of (A.5).

## References

- [1] PL Freddolino, AS Arhipov, SB Larson, A McPherson, and K Schulten. *Structure*, 14:437–449, 2006.
- [2] T Daren, L Perera, L Li, and L Pedersen. *Structure*, 7(3):R55–R60, 1999.
- [3] C Niedermeier and P Taven. *J. Chem. Phys.*, 101(1):734, 1994.
- [4] DJ Griffiths. *Introduction to Electrodynamics*. Benjamin Cummings, 3rd edition, 1999.
- [5] JD Jackson. *Classical Electrodynamics*. Wiley, 2nd edition, 1975.
- [6] LM Burko. *Phys. Rev. E*, 65:046618, 2002.
- [7] R Messina. *J. Chem. Phys.*, 117:11062, 2002.
- [8] G Iversen, YI Kharkats, and J Ulstrup. *Molec. Phys.*, 94(2):297–306, 1998.
- [9] Arfken and Weber. *Mathematical Methods for Physicists*. Harcourt, 5th edition, 2001.
- [10] EW Weisstein. Appell hypergeometric functions. <http://mathworld.wolfram.com/AppellHypergeometricFunction.html>.
- [11] L Lewin. *Polylogarithms and Associated Functions*. North Holland, 1981.
- [12] EW Weisstein. Dilogarithm. <http://mathworld.wolfram.com/Dilogarithm.html>.
- [13] EW Weisstein. Trilogarithm. <http://mathworld.wolfram.com/Trilogarithm.html>.
- [14] Wolfram Research. *Mathematica*. Wolfram Research, Inc., Champaign, IL, 5.2 edition, 2005.
- [15] D Cvijovic and J Klinowski. *Proc. Amer. Math. Soc.*, 125(9):2543, 1997.
- [16] B.R. Brooks, R.E. Bruccoleri, D.J. Olafson, D.J. States, S. Swaminathan, and M. Karplus. *J. Comput. Chem.*, 4:187–217, 1983.
- [17] JC Phillips, R Braun, W Wang, J Gumbert, E Tajkhorshid, E Villa, C Chipot, RD Skeel, L Kale, and K Schulten. *J. Comput. Chem.*, 26:1781, 2005.